ASR

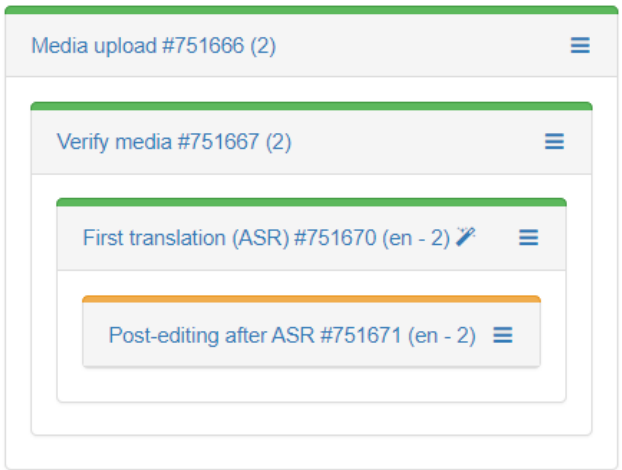
> ASR availability will depend on site settings

> If unsure, you might also want to read up on how parenting works with automated jobs

The "ASR" Delivery method sends the **audio** of the current media for the node to an external service for **Automatic speech recognition**. The result will be parsed into Plint's internal subtitle format and attached to the current node. This means that a subsequent job in *Plint Subtitler* can automatically load this file for further work, such as post-editing. The file could also be automatically converted to a delivery format and used right away, although this is rarely recommended with raw ASR output.

A typical setup:

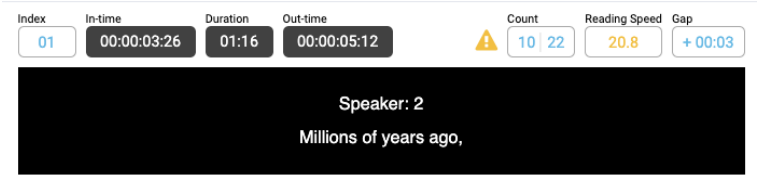
- A project with the **correct source language** specified
- Attach/upload **media** to each part, in typical fashion
- After the media steps (i.e. upload/convert/verify), **add an ASR job**
  - This job can be attached to the part itself, or to a target language
- Follow this with a **online editor job**, for reviewing and adjusting the file



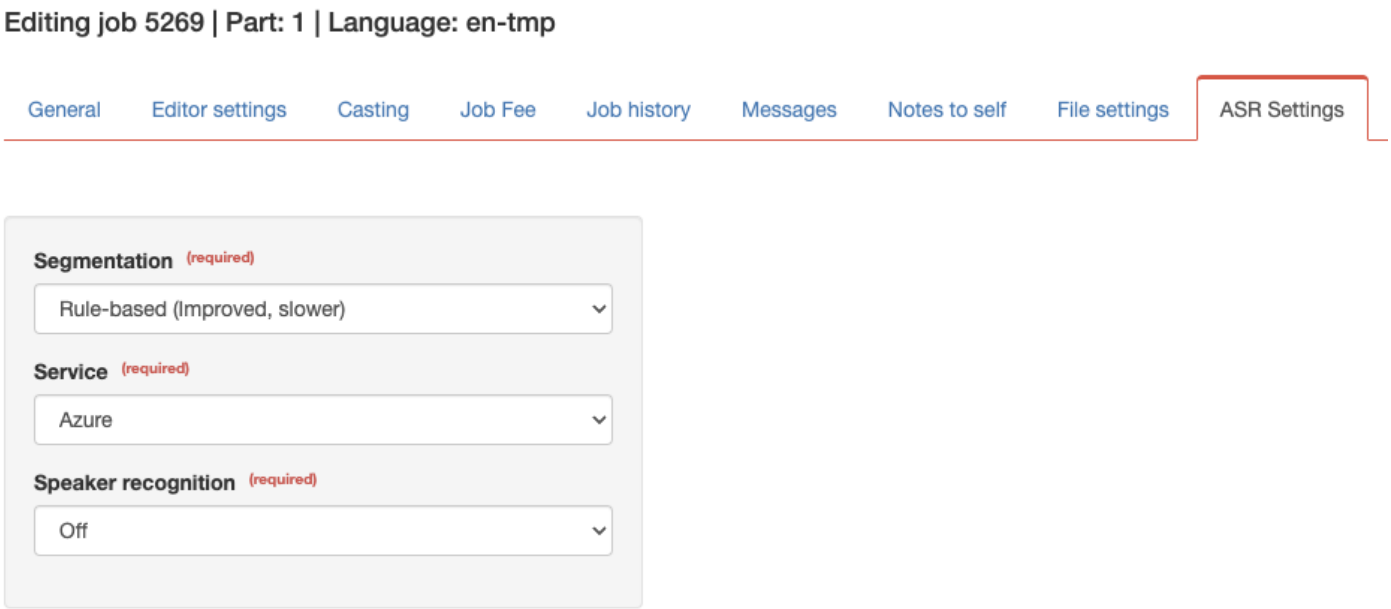
In this example, once the media has been verified and the user changes status, the ASR job will trigger automatically. When this process is completed, the ASR job will automatically change status and the review/post-editing job will become open for work. A subtitle milestone will be stored and can be seen under *Review / Track changes* to enable easy comparison between the original and final versions.

Settings

There are three main settings which can be applied to the ASR job:

- **Service**
  - Depending on what is enabled for your platform, options may include "Azure", "AWS", and others.
- **Segmentation**
  - **Rule-based** segmentation will take into account the Report tool rules matching the job when splitting the text into individual subtitles, basing the subtitle length on the minimum/maximum duration setting.
    - The *Basic* variant will only take the *duration* setting into account when deciding where to end a subtitle. This may require increased post-editing.
    - The *Improved* variant takes longer to process but will reduce the risk of single words ending up at the beginning or end of subtitles.
  - Whereas **Sentence** segmentation will allow longer clips and base the clip length on **punctuation.**
    - This is suitable for non-subtitle work and for when you prefer to do the splitting at the post-editing stage, for example using Plint Subtitler shortcuts.
  - With both types, the *Clip separation* rule matching the project/job will always be applied to ensure that correct spacing is used between subtitles.
- **Speaker recognition**
  - This option will attempt to use the third party service's speaker recognition functionality. The result can be stored either as the first line of the subtitle, or in the Annotation field. The latter field can be either visible, hidden or editable in Plint Subtitler, based on job settings.
  - Please note that with Rule-based segmentation, it is common to get more than one speaker label with the same subtitle. They will all be saved on a single line.
    - Example (Rule-based segmentation, Speaker label stored as part of Subtitle)
    - 

ASR settings tabs



Things to note

- Depending on the segmentation type, certain values will be mandatory in the Report preset in order for the job to run properly. To ensure complete functionality with rule-based segmentation, make sure the Report preset has settings for:
  - Subtitle separation (formerly *Clip separation*)
  - Duration (min and max)
  - Maximum characters per line
- The ASR process can also be run "ad hoc" using the microphone icon in the joblist
- An ASR job set to "Awaiting corrections" will *not* run automatically
  - This status is used to indicate that something went wrong with the process and something will probably need to change before another attempt is made.
- Ensure that the service you select supports the current source language, or the process will fail.
  - For AWS (*Amazon Transcribe*), the list can be found here.
  - For Azure (*Speech to text*), see here.

Revision #19
Created 2021-02-22 12:26:08 UTC by Plint
Updated 2024-04-18 13:41:50 UTC by Plint